McCulloch-Pitts Neuron

medium.com/@manveetdn/notes-on-mp-neuron-padhai-onefourthlabs-course-a-first-course-on-deep-learning-8d2333170d1a

> **Disclaimer: This is notes on " MP Neuron" Lesson (PadhAI onefourthlabs course "A First Course on Deep Learning")**

Six Jars of the Machine Learning:

All the 6-jars put together what is the role of each jar and what we need to do in each stage to built a great Model.

Data: Find out a public source where tons and tons and tones of data sets available like Google, amazon.

Task: In the course we will learn and focus on Supervised learning Classification and Regression.

Learning Algorithm: We are going to focus on stochastic gradient descent, back propagation,Nesterov, RMS prop, momentum based gradient descent.

Evaluation: We are gonna focus on percision,accuracy,recall, top-k, f-score.

We also require:

Linear Algebra ,Probability, Calculus

These are 3 fundamentals areas of mathematics on which you built models.

1. Learning Algorithm that you do is are based on principles and theorems of Calculus [**Maxima, Taylor, gradient, Chain-rule**]
2. We also require Probability we use are **likelyhood, cross-Entropy,KL-Divergence, didtributions**.

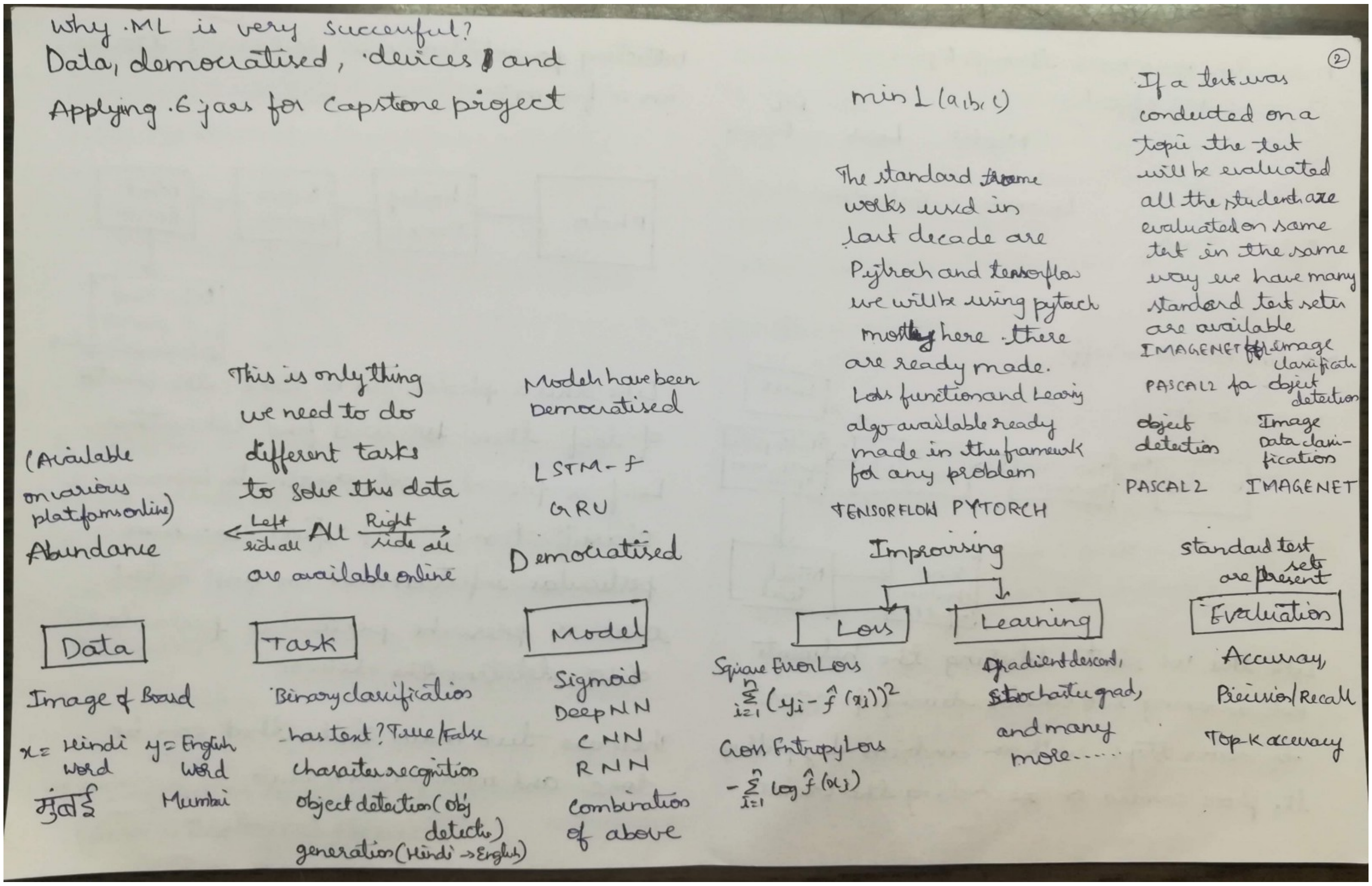
Model:

Some of the models we cover in the couse are f(Wx+b)

here W is matrix R^(mxn), x is vector of R^(1xn).

1. We **use Linear Algebra** like **Vectors, matrix, dot-product, orthogonal vectors and orthogonal matrices ,etc**.

> **We all require above all at one place to solve and built areal time Deep Learning model.**



2

Why ML is very succeful?

Data, democratised, devices and

Applying 6 jars for Capstone project

(Available on various platforms online)

Abundance

Data

Image of Board

x = Hindi word

y = English word

मुंबई

Mumbai

This is only thing we need to do different tasks to solve this data

$\xleftarrow[\text{side all}]{\text{Left}}$ All $\xrightarrow[\text{side all}]{\text{Right}}$

are available online

Task

Binary classification

- has text? True/False

character recognition

object detection (obj detectn)

generation (Hindi → English)

Models have been Democratised

LSTM-f

GRU

Democratised

Model

Sigmoid

Deep NN

CNN

RNN

Combination of above

min $L(a, b, c)$

The standard frameworks used in last decade are Pytorch and tensorflow we will be using pytorch mostly here. there are ready made. Loss function and Learning algo available ready made in this framework for any problem

TENSORFLOW PYTORCH

Improvising

Loss

Square Error Loss

$\sum_{i=1}^{n}(y_i - \hat{f}(x_i))^2$

Cross Entropy Loss

$-\sum_{i=1}^{n} \log \hat{f}(x_i)$

Learning

Gradient descent,

Stochastic grad,

and many more.....

If a test was conducted on a topic the test will be evaluated all the students are evaluated on same test in the same way we have many standard test sets are available

IMAGENET for image classificatn

PASCAL2 for object detection

object detection | Image Data classification
PASCAL2 | IMAGENET

standard test sets are present

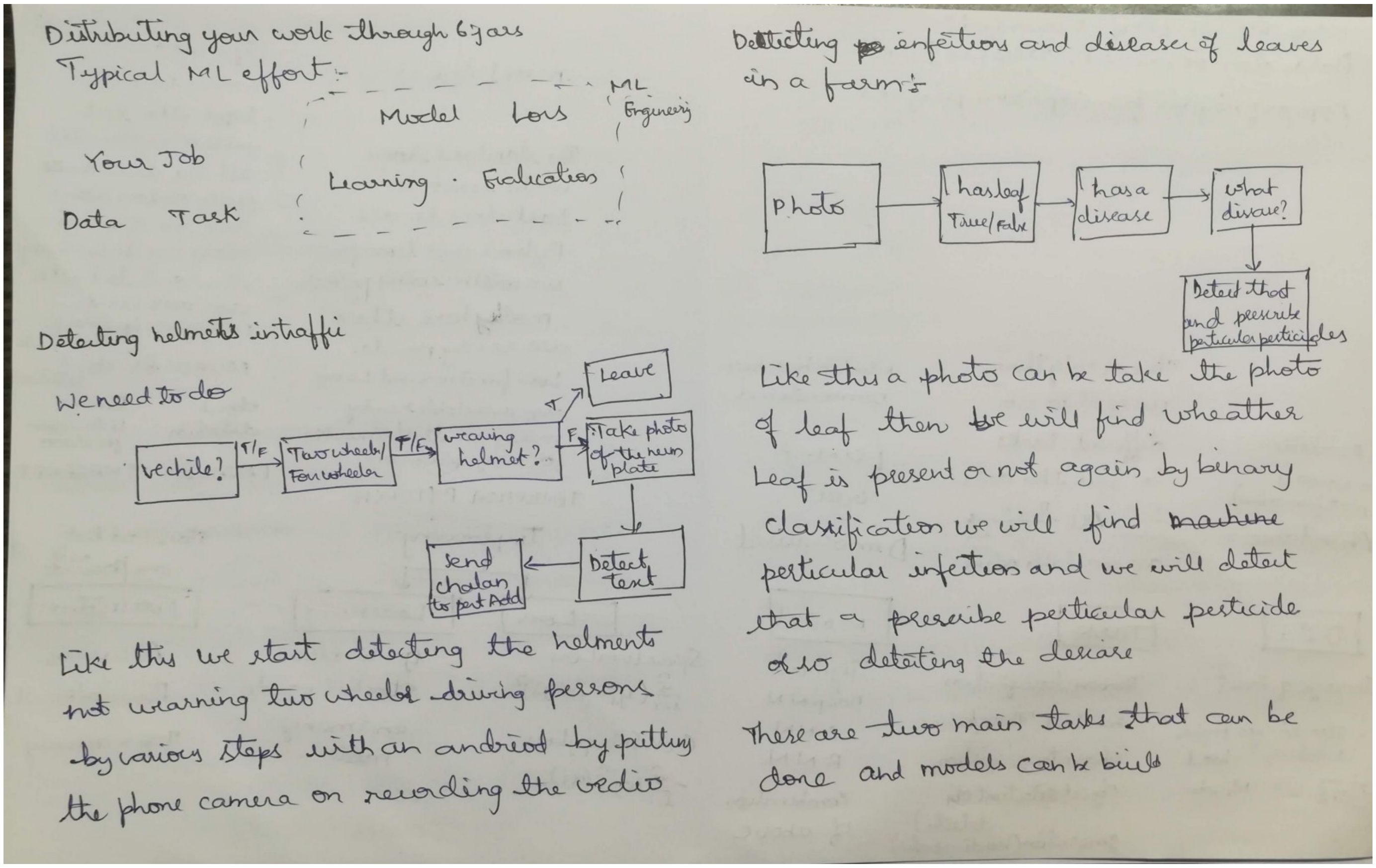
Evaluation

Accuracy,

Precision/Recall

Top-K accuracy

All the 6-jars.



Distributing your work through 6 jars

Typical ML effort :-

ML Engineering
Model
Loss
Your Job
Learning
Evaluations
Data
Task

Detecting helmets in traffic

We need to do

vechile?
T/F
Two wheeler/ Four wheeler
T/F
wearing helmet?
T
Leave
F
Take photo of the num plate
Detect text
Send chalan to perst Add

Like this we start detecting the helments not wearning two wheeler driving persons by various steps with an andriod by putting the phone camera on recording the vedio

Detecting infeitions and diseases of leaves in a farm :-

Photo
has leaf True/False
has a disease
what disease?
Detect that and prescribe particular perticides

Like this a photo can be take the photo of leaf then we will find wheather Leaf is present or not again by binary classification we will find ~~machine~~ perticular infeitions and we will detect that a prescribe perticular pesticide also detecting the desease

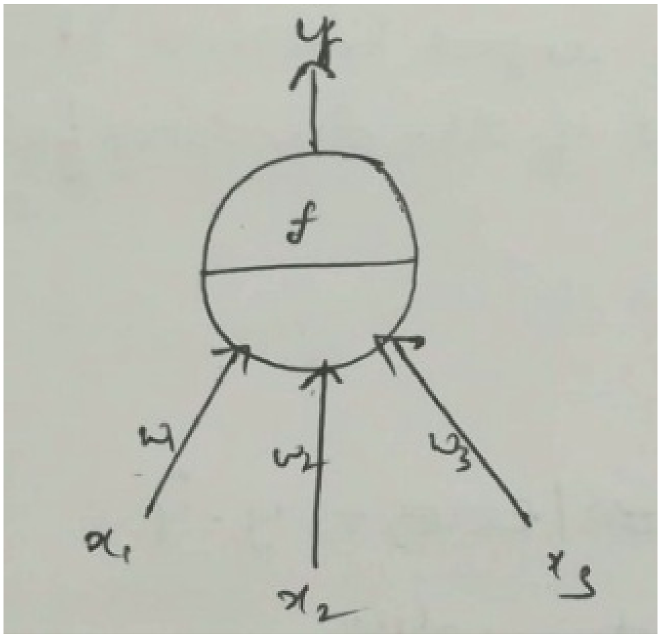
These are two main tasks that can be done and models can be build

Real-time Examples.

McCulloch-Pitts Neuron:

Fundamental block of Deep Learning network is a Artificial Neuron.

Artificial Neuron



It takes input factors x1,x2,x3 and it has added up weights w1,w2,w3 in the above example it applies function and gives a output like **yes/no**, **true/false.**

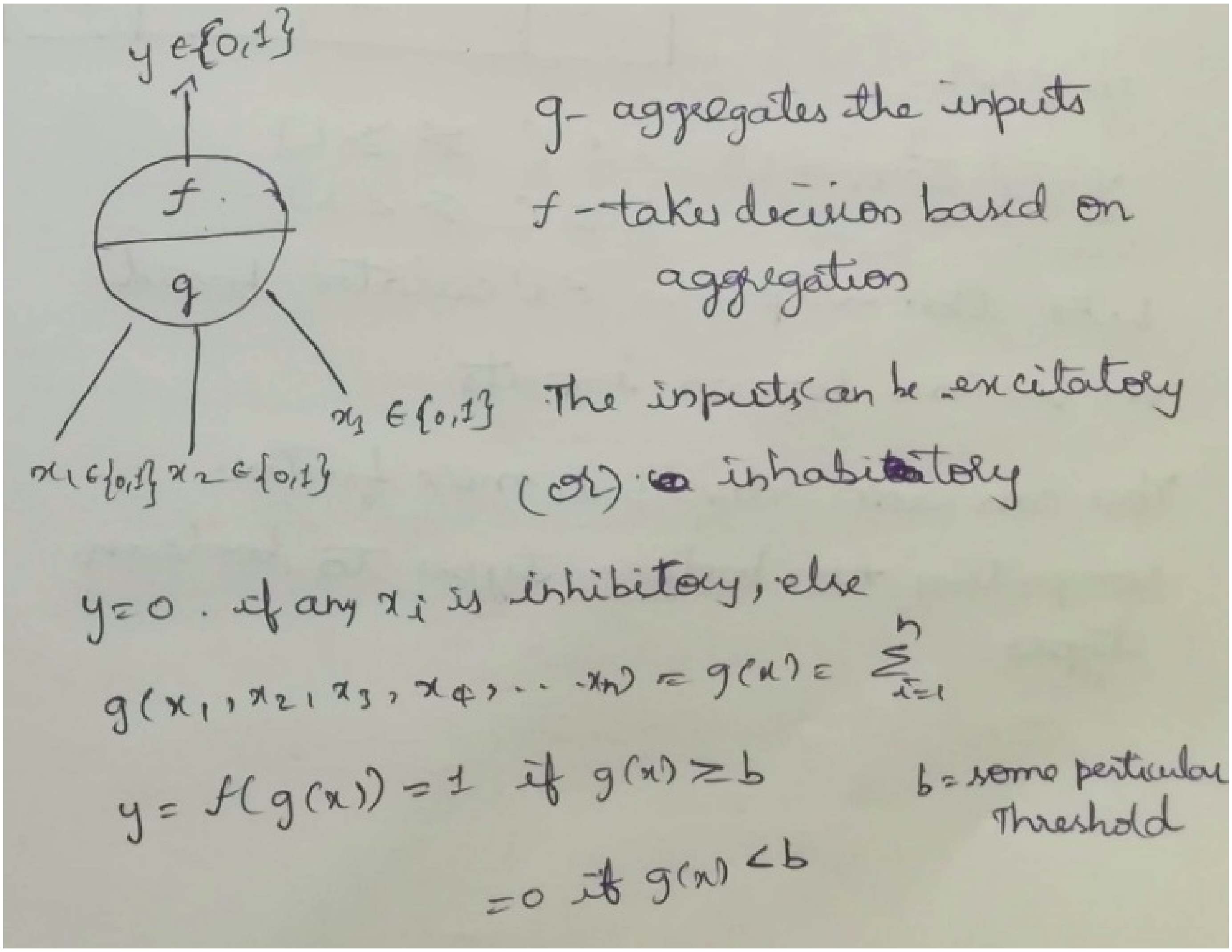
Ex: Going to a Movie.

The various factors which acts as **inputs** for going the movie are **Genere ,Songs,Directors,Actor,Trailer.**

The Inspiration of this Artificial Neuron is from Biological Neurons.

Artifical MP Neuron we proposed by **Watter-Pitts** (A Logician) and **Warren McCulloch**(A Neuro-Scientist**).**

MP Neuron Model:



g- aggregates the inputs

f -- takes decision based on aggregation

The inputs can be excitatory (or) inhabitatory

$y=0$. if any x_i is inhibitory, else

$g(x_1, x_2, x_3, x_4, \dots x_n) = g(x) = \sum_{i=1}^{n}$

$y = f(g(x)) = 1$ if $g(x) \geq b$

b = some perticular Threshold

$= 0$ if $g(x) < b$

MP Neuron Model

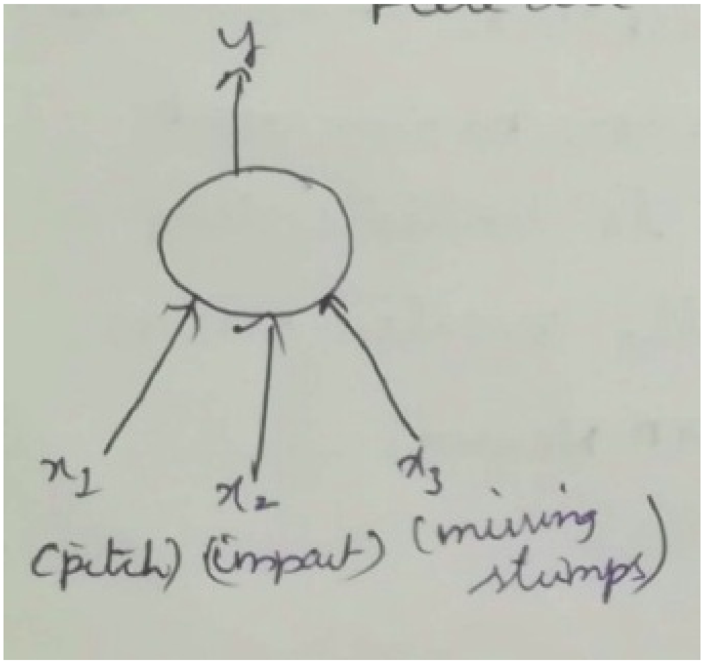
Data and Task:

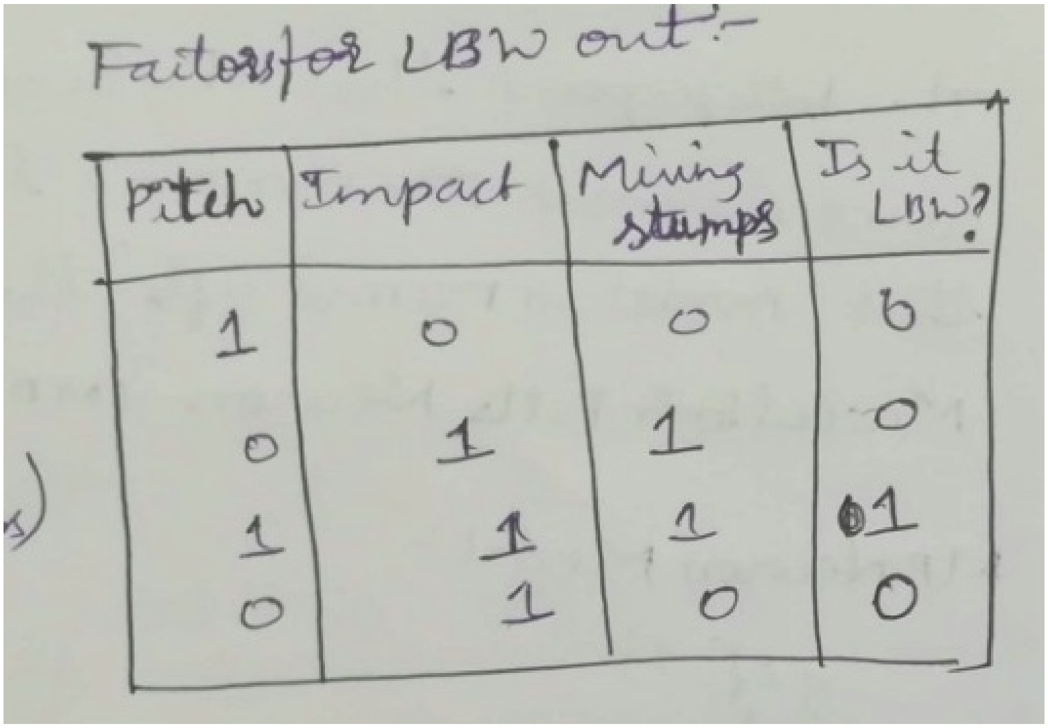
Here, all inputs are boolean values

Data and Task

Here this is basic model MP-Neuron Model taking inputs and giving a output Y which is boolean in nature which is outputs **1 if True** and **0 if False.**

Lets Take the Example of the LBW out in a cricket match.





Factorsfor LBW out:-

Pitch	Impact	Missing stumps	Is it LBW?
1	0	0	0
0	1	1	0
1	1	1	1
0	1	0	0

Factors for an LBW out

W will use the function **y = f(x)** and if the **y = 1 its true or LBW out** and if the final value **y = 0 its False and not Out.** Like that the ouputs is calculated based on the inputs.

$$y = \left(\sum_{i=1}^{3} x_i \geq b \right) \quad \begin{array}{l} = 1 \; (\; \sum \geq b) \\ = 0 \; (\; \sum < b) \end{array}$$

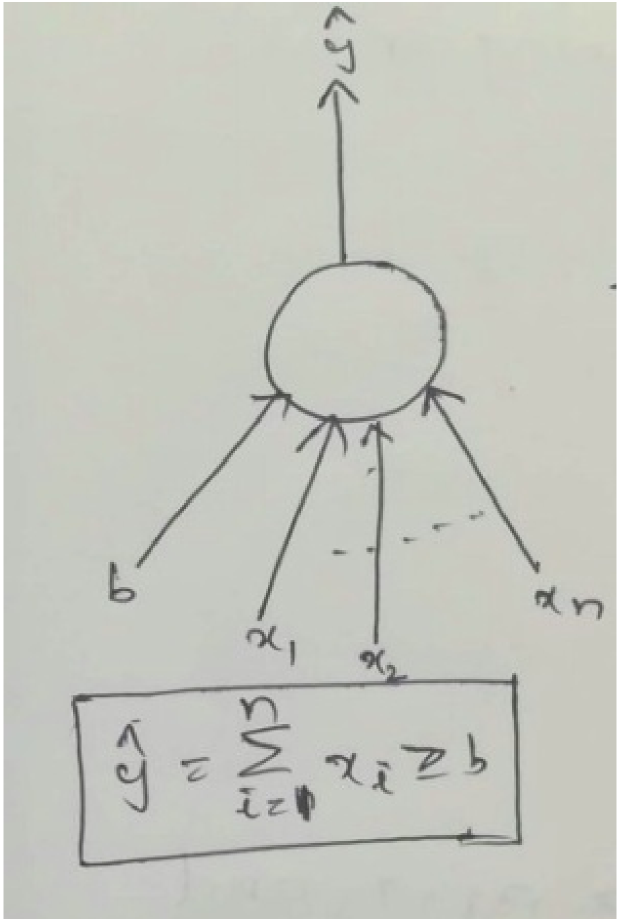
Calculating y for the about LBW-out.

The above is the image how we calculate the y based on the inputs given in the above table of the LBW-out.

You can even described more factors computing non boolean types to boolean types.

Loss Function:

We will also pass threshold as the input because it is also one of the deciding factors.



Loss/Error = [y-(yhat)]

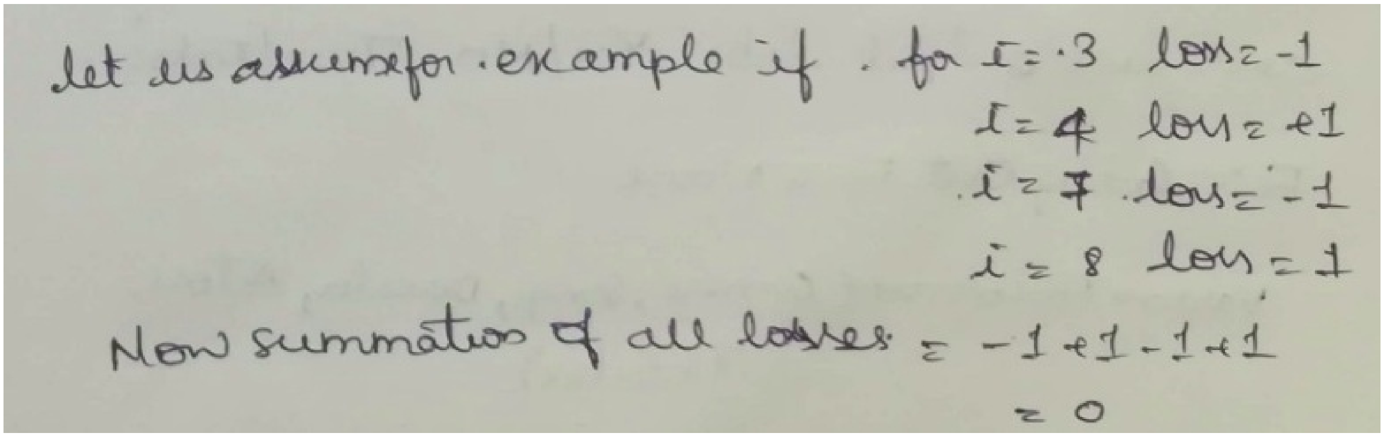
y = true value

(yhat) = predicted value

loss is calculated based on a particular formula, involving the true and the predicted values, its actually the summation of the all the difference for each and every prediction. Like that we will calculate the loss value.

$$\text{loss} = \sum_{i=1}^{n} y_i - \hat{y}_i \quad . \ \text{all loss for all predictions}$$

(Firstly proposed) Loss formula.



let us assumefor example if for $i = 3$ loss $= -1$

$i = 4$ loss $= +1$

$i = 7$ loss $= -1$

$i = 8$ loss $= 1$

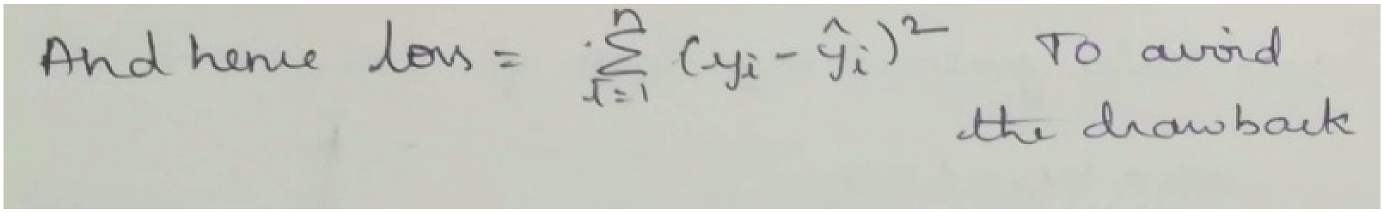
Now summation of all losses $= -1 + 1 - 1 + 1$

$= 0$

Taking an example applying the loss function.

> **We will get the zero as the final output which is a main draw back.There we need a modification in the formulae to predict the accurate loss and overcome the this main draw back.**

Modified Loss Function:



And hence loss $= \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ To avoid the drawback

Modified Loss Function.

We will modify the loss function, based by squaring the difference and summing up all the square of differences.

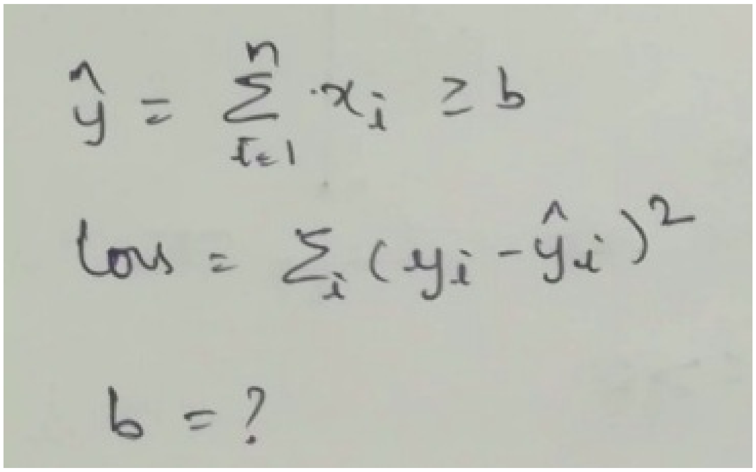
This the formula which overcomes the drawback and suffices the need.

Learning Algorithm:

Learning Algorithm

Now we have the output calculating function to predict the output, loss function and we need the b value thats the bias how will we calculate that.

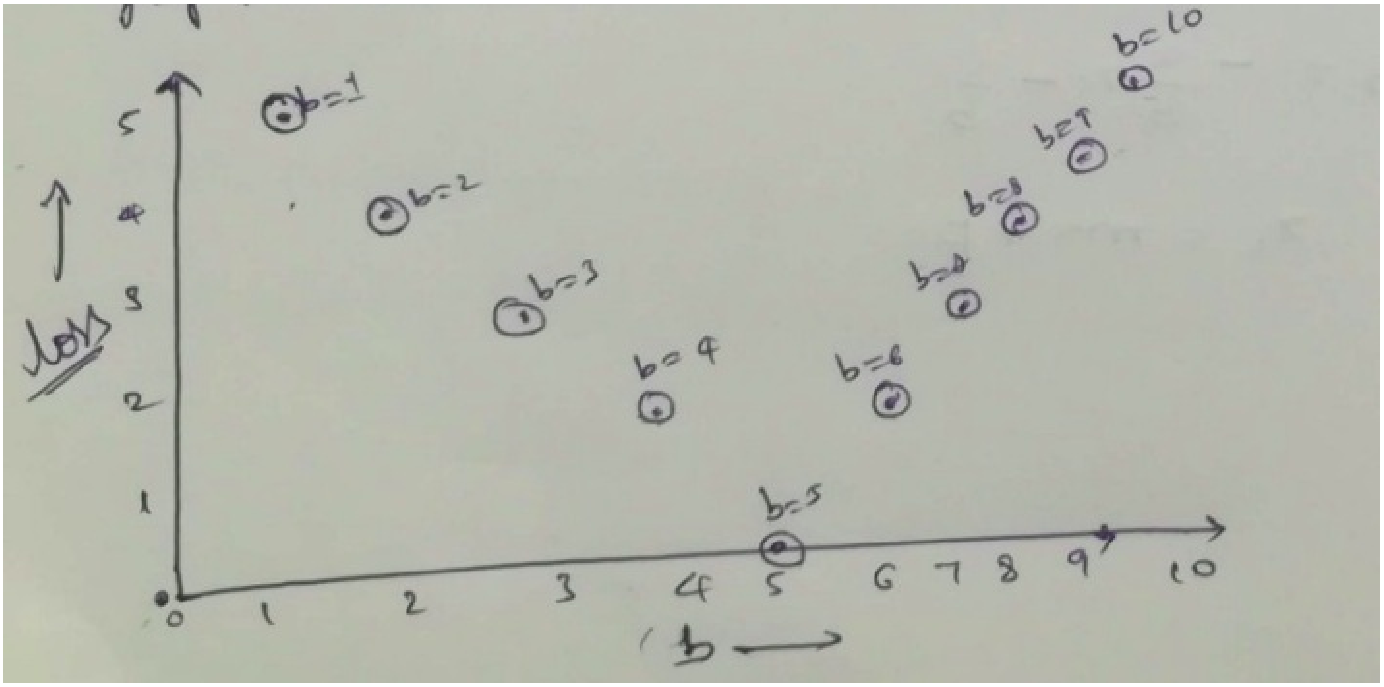
According to a method we take b weight from 1 to 10 and taking each value at a time. we will plot a graph between b and loss and let the graph be as below.



$\hat{y} = \sum_{i=1}^{n} x_i \geq b$

loss $= \sum_i (y_i - \hat{y}_i)^2$

$b = ?$

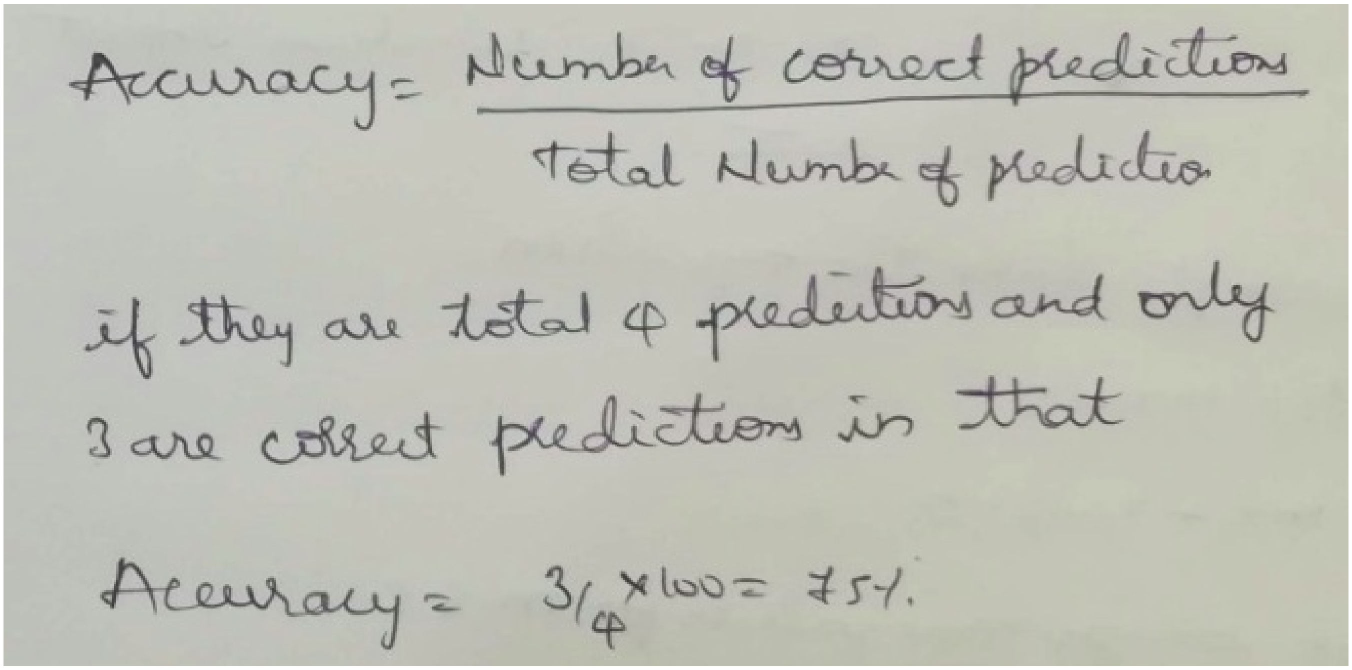


Graph between loss and b (bias).

The output graph is graph as plotted above then **5 is the value of ‘b’ because the loss is zero** at that case like this we will find the b value in the learning algorithm part.

Evaluation:

Accuracy is calculated by the particular formula called the **Accuracy-Formulae.**The ratio of the Number of correction predictions to the total Number of predictions.



$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total Number of predictio}}$$

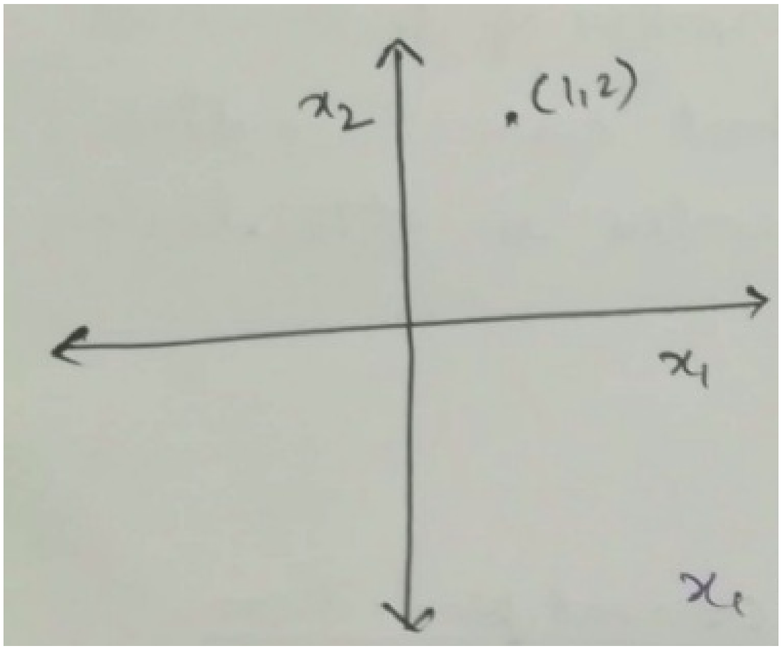
if they are total 4 predictions and only 3 are correct predictions in that

$$\text{Accuracy} = 3/4 \times 100 = 75\%$$

Evaluation Phase.

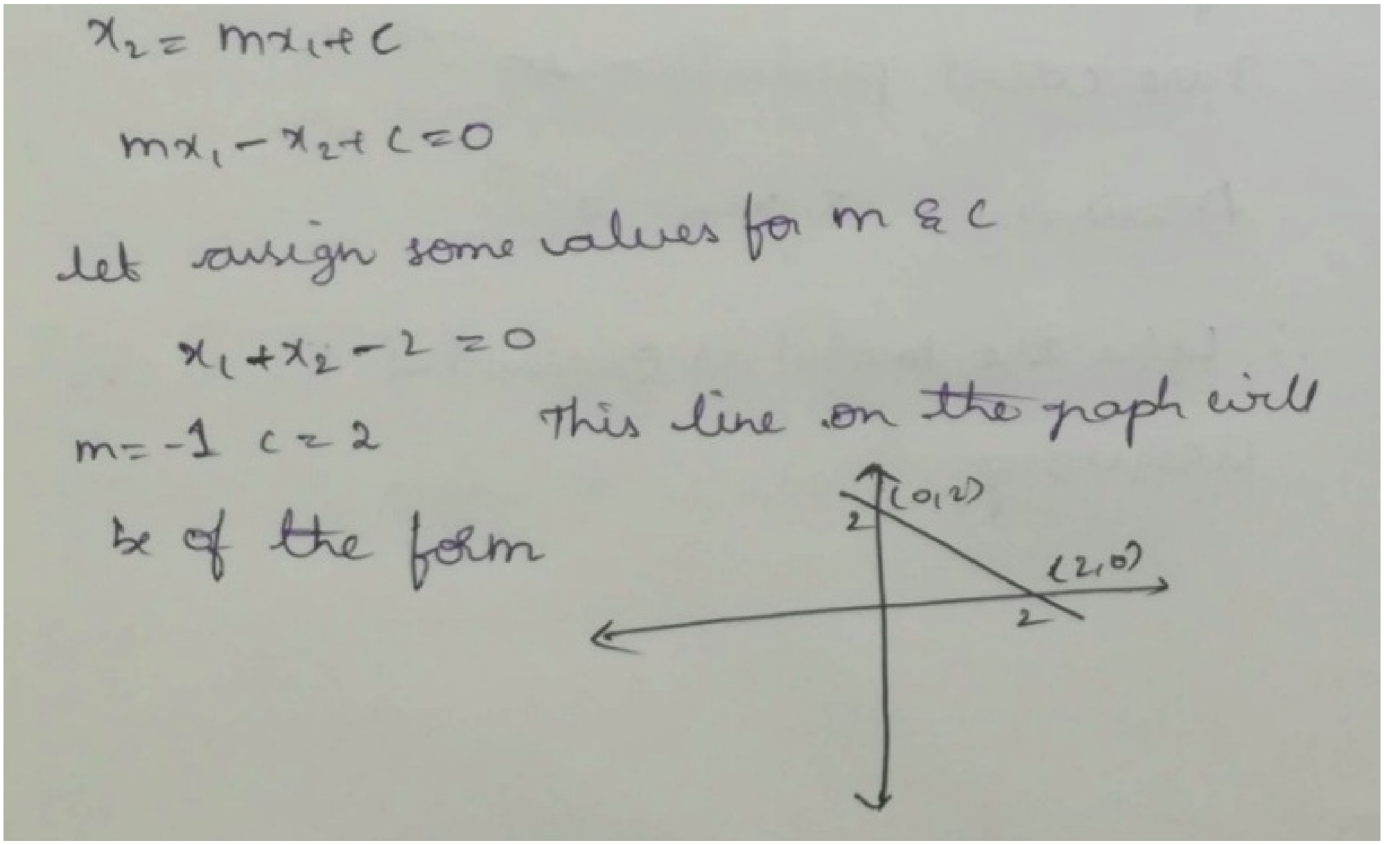
> **Like that we will evaluate the model based on the accuracy value.**

Geometric Interpretation:



Here**, y is the output** that’s why we will not plot between y and x we will plot the graph always between **x1 and x2 which is input data.**

for example take the equation



$x_2 = mx_1 + c$

$mx_1 - x_2 + c = 0$

let assign some values for m & c

$x_1 + x_2 - 2 = 0$

$m = -1$ $c = 2$ This line on the graph will be of the form

Plotting m = -1 and c-= 2

x2 = m*x1 + c

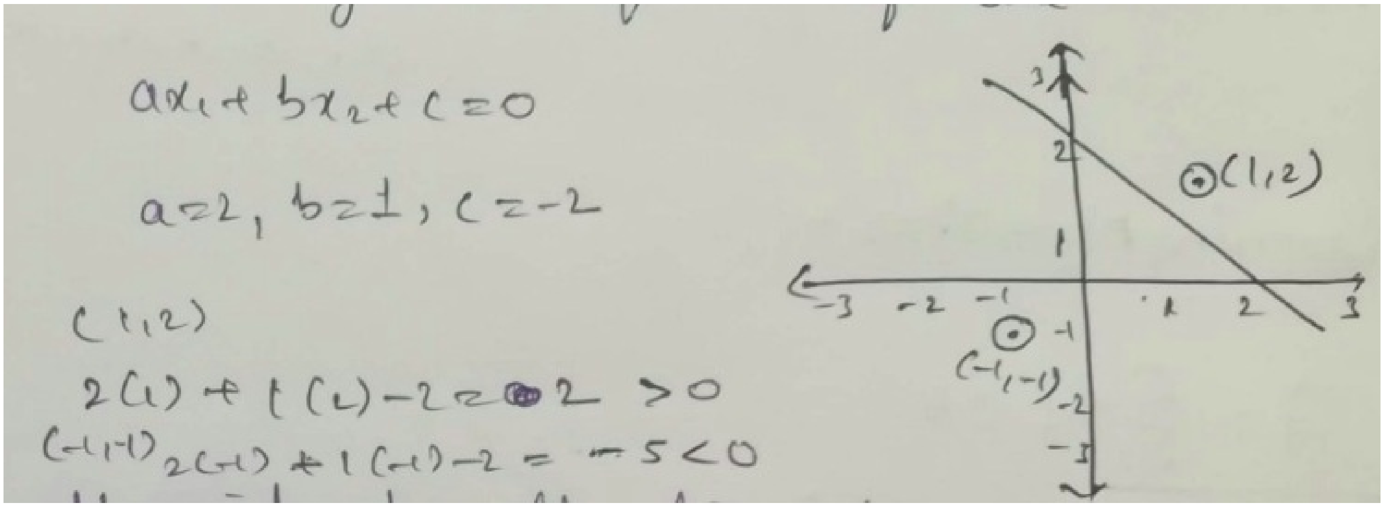
m*x1 — x2+c = 0

Let assign some values for m and c

x1+ x2 -2 = 0, m = -1 and c =2

This on the graph will of the form as beside one.

Lets take the general equation of line



$ax_1 + bx_2 + c = 0$

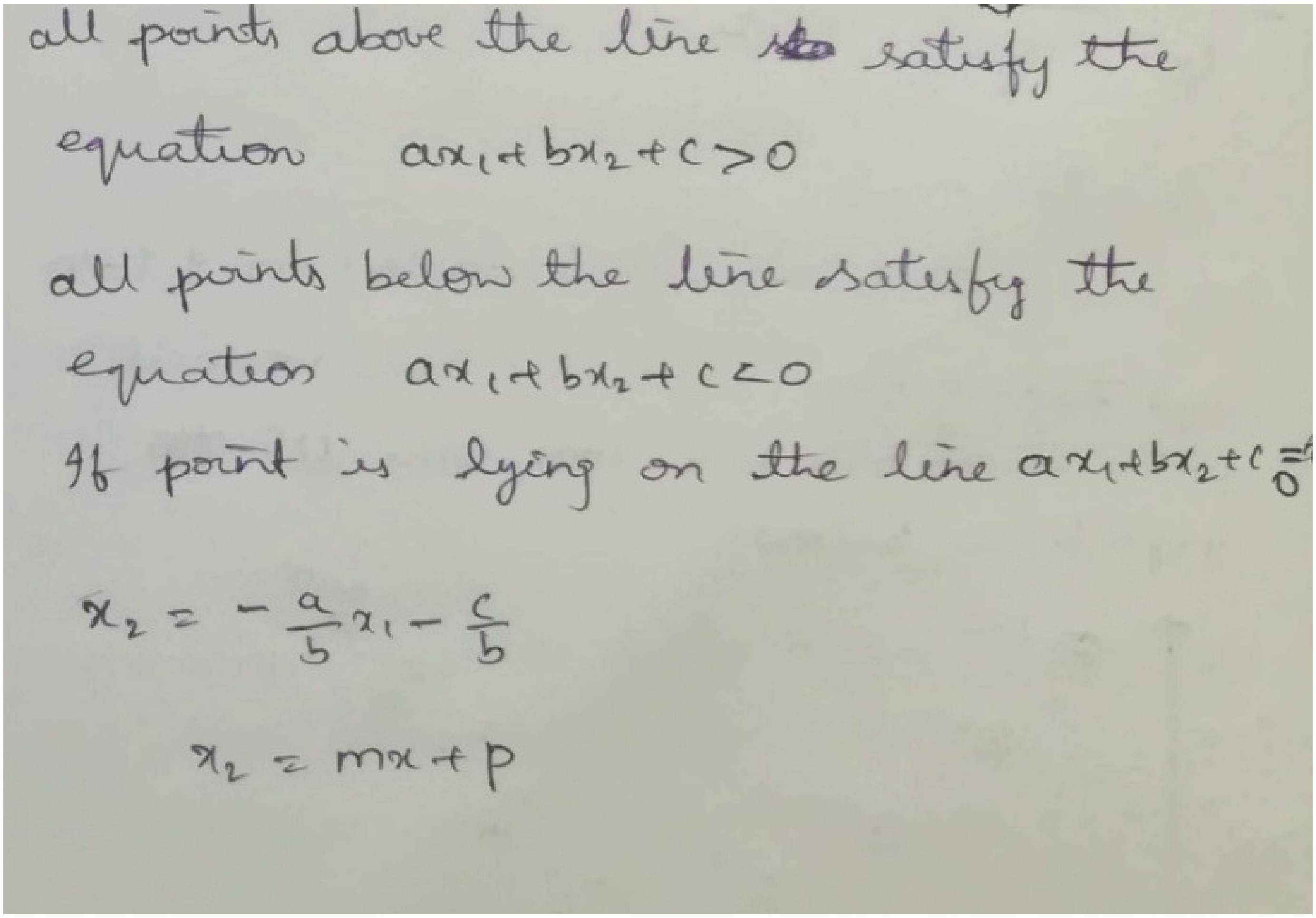
$a = 2, b = 1, c = -2$

(1,2)

$2(1) + 1(2) - 2 = 2 > 0$

(-1,-1) $2(-1) + 1(-1) - 2 = -5 < 0$

a = 2, b = 1 , c = -2 graph.



all points above the line ~~to~~ satisfy the equation $ax_1 + bx_2 + c > 0$

all points below the line satisfy the equation $ax_1 + bx_2 + c < 0$

If point is lying on the line $ax_1 + bx_2 + c = 0$

$$x_2 = -\frac{a}{b}x_1 - \frac{c}{b}$$

$$x_2 = mx + p$$

a*x1+b*x2+c=0

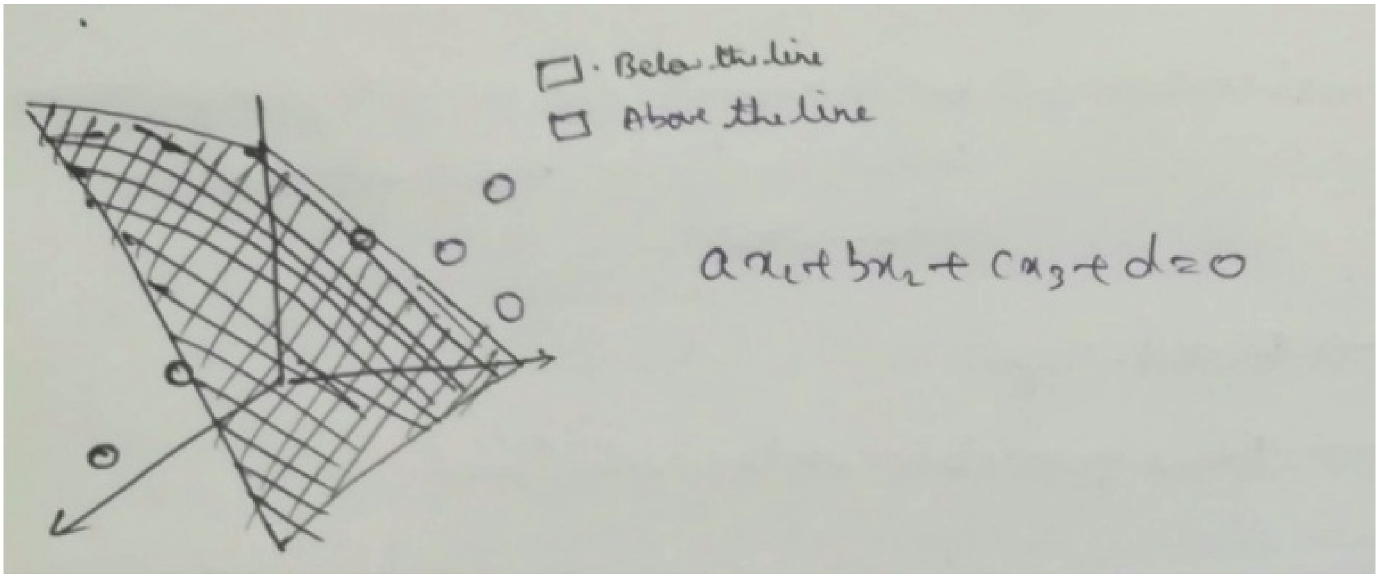
a = 2, b = 1, c = -2

All the points below the line satisf the euation a*x1+b*x2+c<0

If point is lying on the line a*x1+b*x2+c = 0

Like that the points are classification is done.

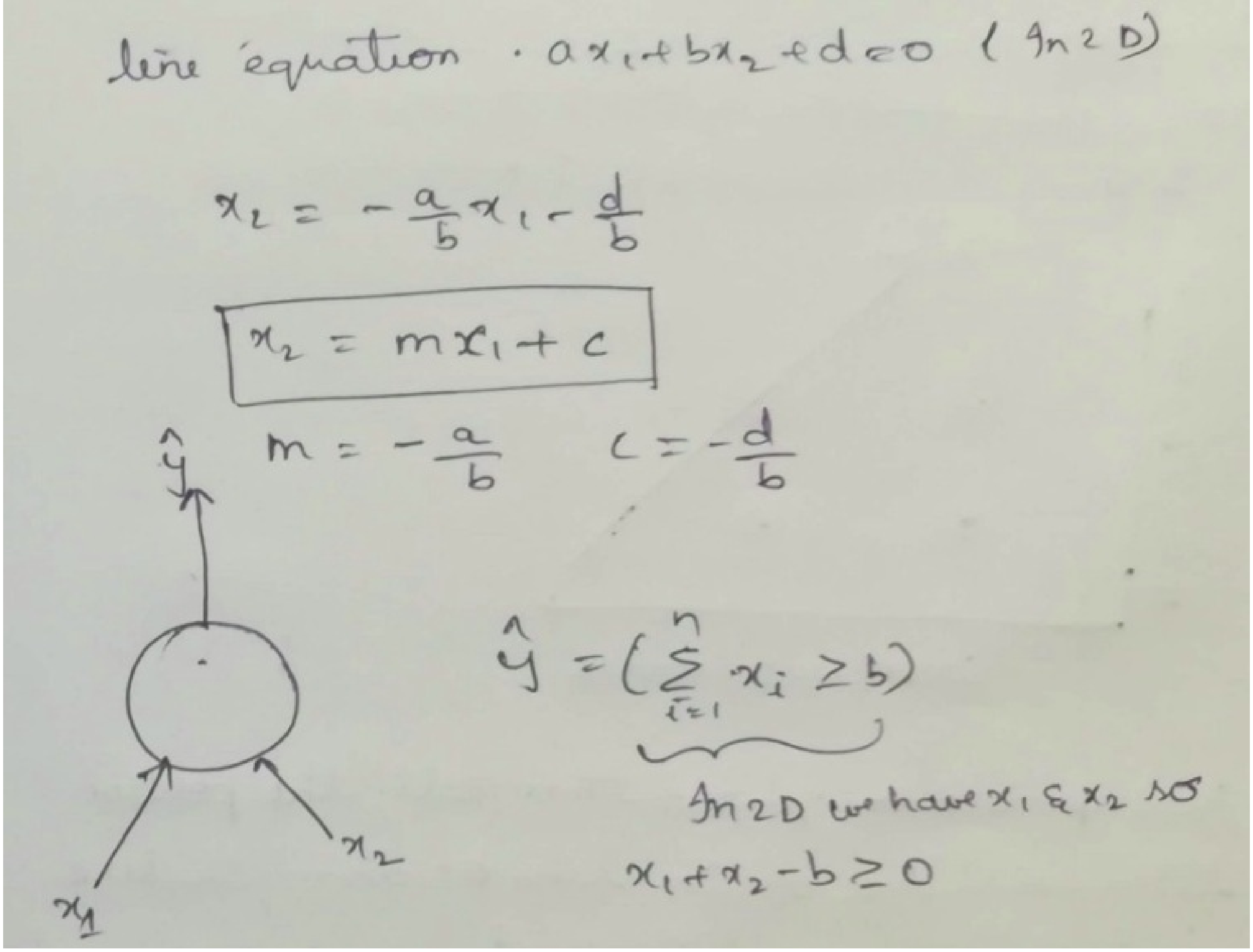
In 3-Dimension:



In 3-D its classified like that.

Geometrical Interpretation:

Line equation a*x1+b*x2+d = 0 (In 2-D)



line equation · $ax_1 + bx_2 + d = 0$ (In 2 D)

$$x_2 = -\frac{a}{b}x_1 - \frac{d}{b}$$

$$\boxed{x_2 = mx_1 + c}$$

$$m = -\frac{a}{b} \qquad c = -\frac{d}{b}$$

$$\hat{y} = (\underbrace{\sum_{i=1}^{n} x_i \geq b})$$

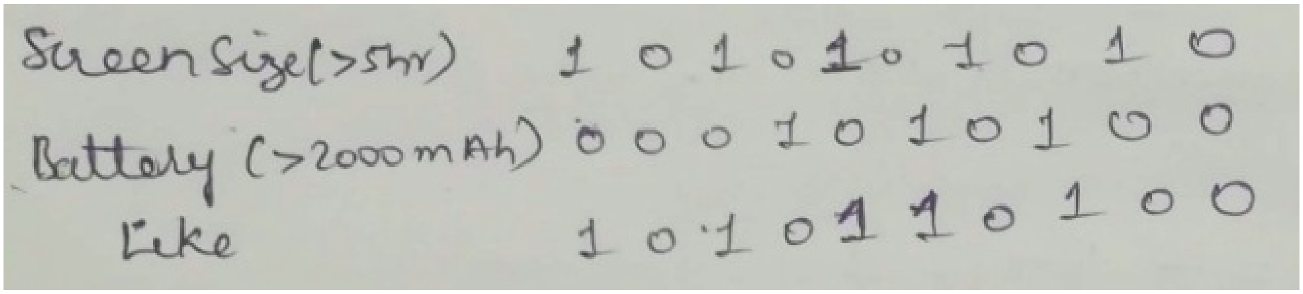
In 2D we have x_1 & x_2 so

$$x_1 + x_2 - b \geq 0$$

MP Neuron does the same thing it divides the whole points into two sets:

1. Set of points above the line(x1+ x2—b ≥ 0)
2. Set of points below the line(x1+ x2—b < 0)

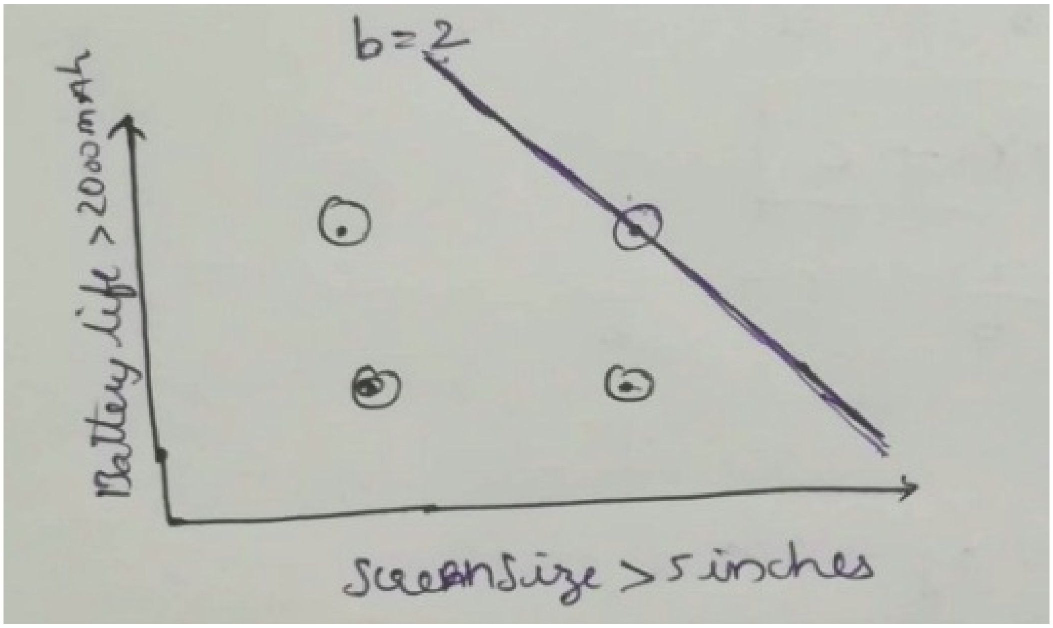
Examples of Cellphones:



Screen Size(>5in)	1	0	1	0	1	0	1	0	1	0
Battery (>2000 mAh)	0	0	0	1	0	1	0	1	0	0
Like	1	0	1	0	1	1	0	1	0	0

Data-set of Cellphones Like/DisLike.

If we take the data set of the having mainly two columns of the **Screen-size** and **Battery life** and need to predict whether users may **Like [1] or DisLike[0].**



Classification.

Here the MP-Neuron shows that all the positive likes will lie on the line on the above the line and another is all the points which are negatives in likes will be negative and below the line.

Dis-advantages of the model:

1. **Linear,** we cannot fit any degree polynomial which be more efficient in other cases.
2. Fixed Slope
3. Few possible intercepts(b's)

This is all about the MP-Neuron Model including Advantages and the Dis -Advantages.

This is a small try ,uploading the notes . I believe in "**Sharing knowledge is that best way of developing skills**".Comments will be appreciated. Even small edits can be suggested.

> **Each Applause will be a great encouragement.**

Do follow my medium for more updates……